

# Associative memory neural network with low temporal spiking rates

DANIEL J. AMIT\*<sup>†</sup> AND A. TREVES<sup>‡</sup>

\*Istituto Nazionale di Fisica Nucleare, Dipartimento di Fisica, Università di Roma 'La Sapienza,' 00185 Rome, Italy; and <sup>†</sup>Racah Institute of Physics, Hebrew University, 91904 Jerusalem, Israel

Communicated by Edoardo Amaldi, May 8, 1989 (received for review November 17, 1988)

**ABSTRACT** We describe a modified attractor neural network in which neuronal dynamics takes place on a time scale of the absolute refractory period but the mean temporal firing rate of any neuron in the network is lower by an arbitrary factor that characterizes the strength of the effective inhibition. It operates by encoding information on the excitatory neurons only and assuming the inhibitory neurons to be faster and to inhibit the excitatory ones by an effective postsynaptic potential that is expressed in terms of the activity of the excitatory neurons themselves. Retrieval is identified as a nonergodic behavior of the network whose consecutive states have a significantly enhanced activity rate for the neurons that should be active in a stored pattern and a reduced activity rate for the neurons that are inactive in the memorized pattern. In contrast to the Hopfield model the network operates away from fixed points and under the strong influence of noise. As a consequence, of the neurons that should be active in a pattern, only a small fraction is active in any given time cycle and those are randomly distributed, leading to reduced temporal rates. We argue that this model brings neural network models much closer to biological reality. We present the results of detailed analysis of the model as well as simulations.

## Section 1. Introduction

Increased activity in the study of properties of attractor neural networks (1–3) has brought about increased pressure for more direct contact with experiment.<sup>§</sup> Such contact may be envisaged on the neurophysiological level (4) or alternatively on the psychological–computational level (5–7). Here we describe a reformulation of the Hopfield model (1) in an attempt to narrow the gap between model attractor neural networks and neurophysiological reality. It is gratifying that despite the richer context captured by the present version its effective dynamics is still quite simple and given to analysis by minor extensions of previously developed techniques (8).

A few of the main “biological” objections to the standard model can be paraphrased as follows.

1. Observed firing rates of neurons in the cortex are very significantly lower than implied by the model. This holds *both* for spatial and for temporal rates.
2. Excitation and inhibition in the cortex (*i*) are affected by different sets of neurons, (*ii*) have different physiological characteristics, and (*iii*) have qualitatively different roles.
3. The synaptic matrix in the models has an awkward feature in that pairs of neurons that are inactive in a memorized pattern have an enhanced excitatory synaptic weight.
4. There is an artificial symmetry between firing and nonfiring neurons in the dynamics of the network.

Some aspects of the above list have been dealt with previously. Low *spatial* firing rates, the fact that usually only a small fraction of the neurons emit bursts, have found various solutions (9–12). But little or no progress has been made on the temporal side. Excitation and inhibition have been separated into different classes of neurons in the network, by diluting the synaptic network of the Hopfield model (13), but their roles have been kept equivalent. Some early studies of neural networks (14) avoid the artificial feature of excitatory connections associated with correlated *inactivity* in learned patterns. They do not, however, opt for an attractor interpretation of their function but are single layer classifiers. Finally, some attention has recently been paid to the distinction between firing and nonfiring neurons (11, 12). The emphasis has been on the marked enhancement in storage capacity that such representations provide. Yet they operate with functionally equivalent excitatory and inhibitory synapses.

The most apparent flaw in the above critical list—the one most directly observable—is that of the temporal spike rates. The basic model (1) discretizes time in units of the order of the *absolute refractory period*, which is 2–3 ms. Retrieval is identified with the arrival of the network at an attractor and the interpretation is that some set of neurons, corresponding to a memorized pattern, emit a spike in every time slice and the other neurons in the network remain essentially quiescent. Alternatively, if the interpretation is in terms of firing rate, the attractors represent states of the attractor neural network in which the same class of neuron fires at maximal frequency and the others fire at very low frequency. Thus the arrival at an attractor, the elementary cognitive event, should be identifiable by the appearance of rates of 300–500 s<sup>−1</sup>, which could be easily observable even by single electrode probes.

Empirically one observes in associative parts of the cortex that when rates actually rise significantly above the spontaneous rate of single spikes per second, in the course of an associative event, they reach 100 s<sup>−1</sup> in some areas (15–17) and 30–40 s<sup>−1</sup> in others (18). If the attractor interpretation is to be sustained the gap of a factor 5–10 in rate must be bridged. Attempts to account for this gap by modifications of the neural response functions lead to rather implausible characteristics of neurons. The model described here is an attempt to account for this gap in terms of collective network properties by taking into account the whole spectrum of features mentioned in the critical list. We have not tried to find a minimal set of modifications since all these modifications seemed desirable.

Coming back to item 2 above, we observe that the separation of inhibition and excitation into separate groups of neurons goes beyond the neurochemical Dale principle (19).

<sup>†</sup>On leave from: Racah Institute of Physics, Hebrew University, 91904 Jerusalem, Israel.

<sup>§</sup>This was strongly impressed on us by the biologists who participated with us in a full-year program at the Institute for Advanced Studies in Jerusalem during 1987–1988.

Usually, pyramidal cells, which are excitatory, predominate in number. Moreover, they are the ones that communicate from one region to another, while inhibitory (stellate) neurons act locally. One might conjecture that information is represented only on the first and that the latter have a subservient role of keeping firing rates where they belong. Second, it is often conjectured that plasticity is most likely to take place around the spines of excitatory neurons, which suggests that most of the synaptic structure should be in the excitatory-excitatory connections. Third, the action of an inhibitory neuron at a synapse differs from that of an excitatory one in that it is much stronger. Sometimes (20) it is described in terms of shunting, though it may be effectively described by particularly high synaptic efficacies.<sup>†</sup> Finally, the local nature of the distribution and communication of inhibitory neurons may indicate that they may learn about the distribution of activity among the excitatory neurons faster than the excitatory ones do and may react on the excitatory ones, in an inhibitory fashion, within a typical time interval in which the excitatory neurons are finding out about the distribution of their own activity.

## Section 2. Modified Attractor Neural Network

**2.1. The Excitatory-Excitatory Network.** In light of the above comments we proceed as follows. The network will consist of  $N$  excitatory neurons and the memorized patterns will be  $p$  activity distributions of these  $N$  neurons. Time will be discretized, as usual, in units of the *absolute refractory period* and a state of the network will be described as an  $N$ -bit word of zeros and ones. A 1 at neuron  $i$  will be interpreted as a spike having been emitted in the corresponding time slice. The patterns will also be words of  $N$  (0, 1)s. The different bits in a pattern are random and each bit  $\eta_i^\mu$ , the state of the  $i$ th neuron in the  $\mu$ th pattern is selected with the distribution  $P(\eta_i^\mu = 1) = a$ , where  $a$  ( $0 \leq a \leq 1$ ) is a parameter that represents the *spatial* mean activity in a retrieval state. As we shall see below, the instantaneous spatial rates will be much lower. The synaptic matrix connecting the excitatory neurons  $i$  and  $j$  will be more closely Hebbian—i.e.,

$$J_{ij} = \frac{1}{Na^2} \sum_{\mu=1}^p \eta_i^\mu \eta_j^\mu, \quad [1]$$

which is composed of contributions of zeroes and ones only and avoids the non-Hebbian feature of item 3. The excitatory postsynaptic potential at neuron  $i$  will be

$$h_i^E = \sum_{j \neq i}^N J_{ij} V_j. \quad [2]$$

**2.2. The Inhibition.** The influence of the inhibitory neurons will be incorporated in an effective way, expressed in terms of the activity of the excitatory ones. If inhibitory neurons can react faster, their effect on the potential of the excitatory ones can be expressed in terms of the activity variables of the excitatory neurons themselves. We will therefore postulate that the inhibitory contribution to the postsynaptic potential of excitatory neuron  $i$  is given as

$$h_i^I = \left( \sum_{\lambda=1}^p \eta_i^\lambda \right) F \left( \sum_{\mu=1}^p \sum_{j=1}^N \eta_j^\mu V_j \right). \quad [3]$$

The first factor is a constant, dependent only on the receiving neurons. Its particular structure is chosen to facilitate our analysis. The function  $F$  represents the mean activity of the

inhibitory neurons and depends on a weighted average of the spiking activity in the excitatory ones.

The function  $F$  is a nonlinear function of its variable. When it is expanded in a power series one observes that the constant term is a threshold and the first-order term acts as an inhibitory contribution to the field of Eq. 2. The first unusual term is the quadratic one. We shall therefore focus our discussion on the effects of this term. The total postsynaptic potential arriving at an excitatory neuron following a given activity distribution among them will then be

$$h_i = \sum_{j \neq i}^N \frac{1}{Na^2} \sum_{\mu=1}^p \eta_i^\mu \eta_j^\mu V_j - \frac{1}{\nu N^2 p^2 a^3} \sum_{\lambda=1}^p \eta_i^\lambda \left( \sum_{\mu=1}^p \sum_{j=1}^N \eta_j^\mu V_j \right)^2, \quad [4]$$

where the coefficient of the quadratic term has been written in a particular normalized way and the new free parameter is  $\nu$ . It will play the central role in what follows. The negative sign has been chosen to make explicit the inhibitory nature of this contribution. Note that the inhibitory term looks like a triadic synaptic connection, for which the potential received by neuron  $i$  depends on the correlation of the activities of pairs of other neurons. We do not imply the necessity of such connections. They arise out of the indirect nonlinear effect of the faster inhibition. The expectation that inhibitory effects are stronger will allow us to obtain low firing rates which require low values of  $\nu$ , see below. The detailed synaptic structure may appear somewhat contrived. This is intended to facilitate analysis. Since, as we shall see below, the network will preferably function under rather noisy conditions, *random* changes in the synapses should not be of great import.

**2.3. The Dynamics and the Relevant Parameters.** The prescription Eq. 4 has to be supplemented with a dynamical rule to complete the description of the network. This will be noisy threshold dynamics—i.e.,

$$\Pr(V_i(t+1) = 1) = [\exp(-\beta h_i) + 1]^{-1} \\ [= 1 - \Pr(V_i = 0)], \quad [5]$$

with  $h_i$  computed on the basis of the configuration at time  $t$  (21). The parameter  $T = \beta^{-1}$  is a measure of the amount of noise. This dynamical system produces a stochastic trajectory of network states from any initial distribution of spiking activity of the excitatory neurons. The nonergodic behavior of such a network is the potential basis for associative memory, as in the standard model (see, e.g., ref. 3).

To classify the nonergodic behavior we define a set of parameters (order parameters) that characterize the cooperative properties of the network. These will be the overlaps (retrieval quality) with the stored patterns

$$x^\mu = \frac{1}{Na} \sum_{i=1}^N \eta_i^\mu \langle V_i \rangle, \quad [6]$$

where  $\langle \rangle$  indicates a temporal average, the total activity in the network,

$$x = \frac{1}{N} \sum_{i=1}^N \langle V_i \rangle, \quad [7]$$

and finally

$$y = \frac{1}{N} \sum_{i=1}^N \langle V_i \rangle^2. \quad [8]$$

The overlap parameters are somewhat different from those used for the Hopfield model. They measure the similarity of a state to a pattern only among the neurons that are expected

<sup>†</sup>We owe this observation to H. Sompolinsky.

to be active in the stored pattern. In fact  $0 \leq x^\mu \leq 1$ . As for  $y$ , it measures the degree of “freezing”—i.e.,  $y \approx x$  implies that the same neurons fire in consecutive time steps while  $y \approx x^2$  implies that the mean activity level is produced by totally uncorrelated neurons at different time steps.

### Section 3. The Scenario

Low temporal firing rates are to be expected according to the following scenario: When a pattern (no. 1) is retrieved

1. a fraction  $x^1 \gg x$  of the neurons corresponding to the ones in that pattern will be active at every instant,
2.  $x^\mu \ll x^1$  for  $\mu \neq 1$ ,
3. only a small fraction of the neurons that should be active in a pattern is indeed active at any moment; i.e.,  $x^1 \ll 1$ , and
4. the fraction of active neurons in each of the two groups, corresponding to the ones and the zeros in pattern 1, respectively, are uncorrelated at different instants in time. This can be expressed in terms of  $y$ , as the condition

$$y = a(x^1)^2 + (1 - a)(\bar{x}^1)^2. \quad [9]$$

The first two items reflect the requirement that retrieval be unambiguous. They imply that the fraction of active neurons among those that should be quiescent in the retrieved pattern, given by

$$\bar{x}^1 \equiv \frac{1}{(1 - a)N} \sum_{i=1}^N (1 - \eta_i) \langle V_i \rangle = x - \frac{a}{1 - a} (x^1 - x) \quad [10]$$

be much smaller than  $x^1$ . The last two are our way to low rates. Since if the fraction of  $x^1$  neurons that should be active is uncorrelated from one instant to the next, due to noise of course, then the mean temporal rate of these neurons will be simply  $x^1$  times the basic rate. On the other hand, if the same holds in the neurons of the second group, then their temporal rate will be reduced by a factor of  $\bar{x}^1$ , which according to Eq. 10 is much lower.

To compare and contrast this scenario with the performance of a Hopfield network, note first that the essential similarity is in the fact that retrieval is defined as the appearance of a temporal sequence of network activity states, all unambiguously correlated with a stored pattern. That pattern is said to be associatively retrieved by the stimulus. But in a Hopfield network the activity state is essentially the same throughout the sequence, with at most a small fraction of the neurons changing their state from one time step to the next. Here, instead, the state of the network changes completely at each time step, and the only feature subsequent states have in common is their distinctively high correlation with the stored pattern being retrieved. To bring out the full picture coded in the pattern the activity must be averaged over a few mean spike intervals. But this is similar to the duration required for establishing that an attractor has been reached (see, e.g., ref. 3).

This contrast is reflected in our attitude to noise. It is fast noise that produces this ever-changing instantaneous firing configuration, by determining which neurons, among those that have received an adequate input, will actually fire at any given time. If this noise were to decrease, the network would stop functioning properly, not because of spurious attractors (22), but rather because some particularly excitable neurons would “freeze” in their firing state, and thus appear to fire at inordinately high rates. Thus, for example, the question of storage capacity as discussed previously (9, 22) becomes meaningless, since the model will not be operated without

noise. Capacity will be a function of the desired temporal rate and of the noise level within its permissible window.

### Section 4. Performance of the Network

The description of the performance of the network becomes particularly simple when it is assumed to contain a very large number of neurons  $N$  and to store a large number of patterns  $p$ . The ratio  $\alpha = p/N$  will be finite and sometimes very small, in the limit  $N \rightarrow \infty$ .

**4.1 Qualitative Behavior.** At very high noise levels,  $T(= 1/\beta) \gg p$ , the network is ergodic and each neuron has a probability  $1/2$  of firing at any instant. Then,  $x^\mu = x = 1/2$  and  $y = 1/4$ . As the noise decreases the network remains *ergodic*—i.e.,  $x^\mu = x$  and  $y = x^2$ —but the values of these variables begin to decrease. At noise levels of order unity, but above  $T_R$ , the network is ergodic but  $x^\mu = x = \nu$  and  $y = \nu^2$ . The lower  $\nu$ , the stronger the effective inhibition, and the lower the uniformly distributed activity which is reached by the network independently of the initial conditions. This freezing of the mean activities is represented at the top of Fig. 1, which was obtained as a solution of the full equations for the system; it will be rationalized in Section 4.3 below and was derived in a technical account (23).

Following the long inhibition-controlled ergodic region, at  $T = T_R(a, \nu)$  retrieval appears discontinuously. A new noise window opens up bounded from below by  $T = T_C(a, \nu)$  (Fig. 1). In this noise range ergodicity is broken. Following a short transient, and depending on the initial state of the network, one pattern is clearly distinguished. While the overall mean activity is still controlled by the inhibition (i.e.,  $x = \nu$ ),  $x^1$  becomes larger and  $\bar{x}^1$ , much smaller than  $\nu$ . For the other patterns  $x^\mu = \nu$ . That is to say, the memory structure manifests itself by the selection of one pattern for which the activity among the neurons corresponding to ones in the pattern increases at the expense of the activity of neurons corresponding to zeros in the same pattern. The total activity is kept constant. The disordered state remains for a while an

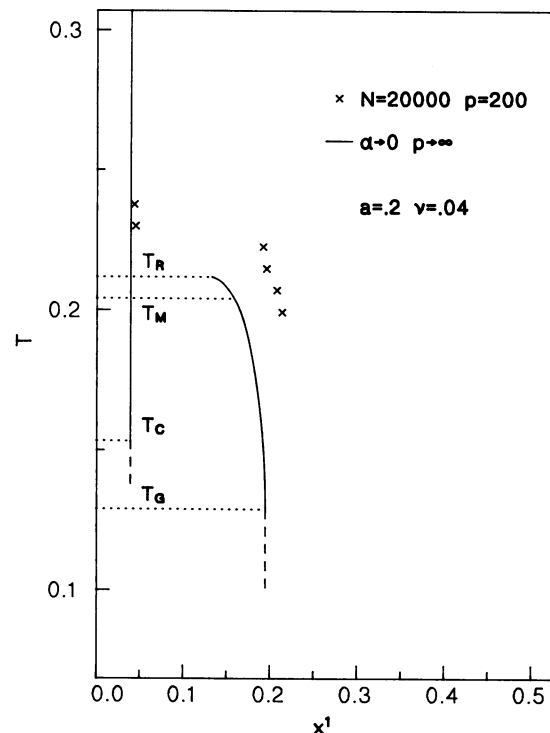


FIG. 1. Overlap with the retrieved pattern,  $x^1$ , as a function of noise level  $T$ . —, Result obtained by solving the mean field equations in the limits indicated; x, sample results of simulations.

attractor, as in the standard model (9). It is destabilized below a noise level  $T_C$  (Fig. 1).

The scenario is fully realized in this noise window if  $\alpha$  is very small. First, we find that Eq. 9 is satisfied, implying that the active neurons fire in an uncorrelated fashion and the rate is reduced by the factor  $x^1$ . If one chooses  $\nu < a$  then one always has  $x^1 < \nu/a$ , which is the asymptote of the  $x - T$  curve in Fig. 1, as  $T \rightarrow 0$ . The initial value of  $x^1$ , as  $T$  goes below  $T_R$  is already significantly greater than  $\nu$  and it increases as the noise is lowered. For  $a = 0.2$  and  $\nu = 0.04$ , for example, we find  $x^1/\nu = 3.5$ . For this value of  $x^1$  one has, by Eq. 10,  $\bar{x}^1 = 0.015$ , which is much smaller than  $x^1$ . Thus we fix the spontaneous rate at the level of  $\bar{x}^1$ , which can be made very low, since it goes to zero as  $x^1$  approaches its asymptotic value of  $\nu/a$ , linearly in  $(\nu/a) - x^1$ . While  $x^1$  never reaches this asymptotic value, because  $T_G$  intervenes as  $T$  is lowered, it does approach rather close, as is shown by the example mentioned above in which, when  $T \approx T_G$ ,  $(\nu/a) - x^1 \approx 0.002$  and  $\bar{x}^1 \approx 0.0005$ . Thus retrieval is unambiguous. For finite small  $\alpha$  this picture persists approximately, with corrections of  $O(\alpha)$  to Eq. 9. Such corrections imply a nonuniform distribution of firing rates in the two groups of neurons whose width is  $O(\sqrt{\alpha})$ . Another bonus of this model is the absence of the doubling of the memory states due to the symmetry between active and passive neurons.

When noise is lowered below  $T_G$  freezing sets in even at very very low values of  $\alpha$ , as is manifested by the violation of Eq. 9. In fact,  $y$  increases above the random value of the right-hand side of Eq. 9. This has two negative effects. Some neurons fire at a rate that is too high. But in addition static errors will appear in retrieval, and freezing means also that some of the 1-neurons in the pattern may never fire, while some of the 0-neurons will fire at high rate. (At  $T_G$  the network undergoes a spin-glass transition in the limit of  $\alpha \rightarrow 0$ .)

**4.2. Simulations.** Simulations confirm the picture emerging from analysis of the mean-field equations. At high levels of noise the network wanders freely among the states, independent of the initial configuration, and all the overlaps with the stored patterns are equal to the mean firing rate. Below a critical noise level ergodicity is broken and the system flows toward the attractor corresponding to the pattern that has the highest overlap with the initial configuration, and its stochastic wandering is restricted to this valley. If the initial configuration had equal overlap with all the patterns, one of them is selected at random. At low noise levels the network progressively freezes into one of the many states in which a small fraction of neurons fires at maximum rate.

As the noise is lowered one observes the much higher sensitivity of the present network (compared to the Hopfield one) to effects of finite size, which are to be expected. For example, the control of the level of activity requires  $p$  to be large, but this conflicts with the requirement that  $p/N$  be small, to reduce the level of slow noise introduced by high loading. [Finite-size effects have been discussed in detail elsewhere (23).] The mean-field equations represent the limit of the infinite system and hence quantitative differences appear when  $p$  and  $N$  are finite. Typical values were  $p \approx 50$ –200,  $N \approx 10,000$ –50,000. In particular the critical noise levels are shifted from their asymptotic values, as are the equilibrium values of  $x$ ,  $x^\mu$ , etc; Eq. 9 is satisfied only approximately. Such effects can be accounted for by extensions of the analysis.

In Fig. 2 we present a sample simulation. Note that if the basic cycle time is 2.5 msec, the mean rate of the active neurons in the case shown is of 80 spikes per sec.

Because of the finite size not all neurons exhibit the mean rate. Some lapse into undesirably high rates. If the model were to describe a realistic neural system such effects would have to be expected. At finite  $\alpha$  there are corrections to Eq. 9, which imply that the firing rates of individual neurons are

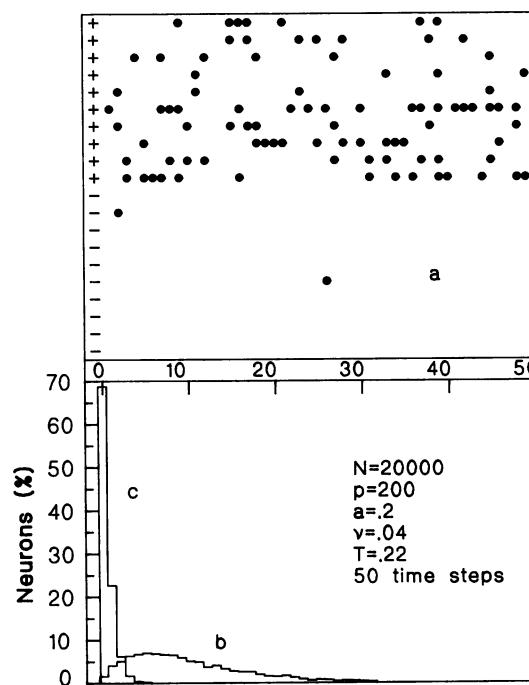


FIG. 2. (a) Firing pattern of 20 randomly chosen neurons out of a network of 20,000 neurons storing 200 patterns, with  $a = 0.2$ ,  $\nu = 0.04$ ,  $T = 0.22$ . •, Spikes at a time proportional to the displacement along the abscissa, whose total length is 50 basic time cycles. The column on the left indicates whether the neuron on the corresponding line is active or passive in the retrieved pattern. (b) Distribution of number of spikes emitted in 50 time cycles by neurons active in the retrieved pattern. (c) Distribution of number of spikes emitted in 50 time cycles by neurons passive in the retrieved pattern.

not very sharply peaked about  $x^1$  and  $\bar{x}^1$ , respectively. This has a rather realistic flavor, provided too-high rates do not become popular. The simulations indicate that the proposed model has a self-correcting mechanism against the appearance of high rates. It turns out that the neurons with excessive activity are those for which  $\sum_\mu \eta_i^\mu < pa$ —i.e., those whose coupling in the inhibitory term (Eq. 3) is too weak. These are also the neurons that, due to fluctuations, participate in fewer patterns. Because it is the weakly coupled neurons that are excessively active it is possible to envisage, in the context of a more comprehensive model that includes learning, a compensatory (Hebbian) effect that, by enhancing the synaptic strength of those neurons that fire more, will in turn reduce their mean activity by means of a stronger inhibition.

An interesting finding on the simulations, which deserves further investigation, is the apparent absence of spurious states that are attractors with a finite overlap with several patterns. We have never encountered such stable states even when the initial configuration has been such a mixture. On the other hand they appear when one takes the inhibitory effective field to depend simply on the mean activity of the excitatory neurons rather than on the weighted averages of Eq. 3. Another feature that emerges in the simulations is the much shorter relaxation time for the stabilization of the average activity of the network as a whole as compared with the time required for the stabilization of the retrieval overlap. In the light of the comments below one can understand this feature as a manifestation of the fact that the patterns are small valleys in an immense basin corresponding to a prescribed mean total activity.

**4.3. Some Qualitative Arguments.** The detailed structure of the synaptic efficacies in the effective field of the excitatory neurons, Eq. 4, was chosen to ensure that the dynamics of the network be governed by an energy function and in the noisy

situation by a free energy. This facilitates the analysis and provides many useful intuitions. The “energy” of a given configuration  $\{V_i\}$  of the excitatory neurons is

$$E\{V\} = -\frac{1}{2N a^2} \sum_{i,j} \sum_{\mu=1}^p \eta_i^\mu \eta_j^\mu V_i V_j + \frac{1}{3\nu N^2 p^2 a^3} \sum_{\mu,\lambda,\sigma} \sum_{i,j,k} \eta_i^\mu \eta_j^\lambda \eta_k^\sigma V_i V_j V_k. \quad [11]$$

Note that even when a network state is totally uncorrelated with the patterns the energy is of order  $p$ . This is due to the fact that our patterns and states are represented by zeros and ones. If  $x^\mu = x$ ,

$$E/N = p \left( -\frac{1}{2} x^2 + \frac{1}{3\nu} x^3 \right). \quad [12]$$

Since  $p$  is large, this quantity dominates the energy landscape, unless the temperature is larger than  $O(p)$  and the entropy takes over imposing  $x = 1/2$ . As  $T$  decreases the influence of the energy increases and the mean activity shifts toward the minimum of Eq. 12, which is at  $x = \nu$ , and where  $E/N = -p\nu^2/6$ .

When a finite number of  $x^\mu$  values deviate from the uniform value the change in the energy is of order unity. It is these changes in the energy that can bring about a differentiation between the patterns. Thus one finds a large range of  $T$  in which the noise is too low to violate the energetic constraint on the overall activity but high enough with respect to the barriers created by different patterns. This is the region  $p \gg T > T_R$  in Fig. 1. Below  $T_R$  the noise is low enough for the detailed structure, inside the large deep well at  $x^\mu = x = \nu$ , to become effective. But since  $\nu < a$  the new expectation value for one of the  $x^\mu$  values can arise in many different ways. In the noise range  $T_R > T > T_G$  the barriers between different states with the same  $x^1$  are smaller than the noise and the network wanders freely in this subspace. This ensures that Eq. 9 is satisfied in the form

$$y = a(x^1)^2 + (1 - a) \left( \frac{\nu}{1 - a} - \frac{a}{1 - a} x^1 \right)^2. \quad [13]$$

Below  $T_G$  the noise is low enough to allow the barriers between the different microscopic realizations of a given value of  $x^1$  to become operative. There spin-glass effects come into play and our discussion stops. Inspecting Fig. 1 one observes some further structure. When retrieval appears it is still a local minimum of the free energy. The paramagnetic phase, with all  $x^\mu = \nu$  is still the global minimum. The retrieval state becomes a global minimum at  $T = T_M$ , above yet another notable noise level  $-T_C = \nu(1 - \nu)(1 - a)/a$ . At

this temperature the paramagnetic, uniform state becomes continuously unstable.

These results can be formally derived from the equations for the minima of the free energy in the replica symmetric approximation, which is expected to be precise since no spin-glass effects appear for small  $\alpha$  and  $T > T_G, T_C$ . The equations, which have been discussed in detail elsewhere (23), are similar to the equations for the standard model (9).

We are indebted to Prof. M. Abeles for impressing the gravity of the rate problem on us and to Profs. H. Sompolinsky, E. Gardner, L. Abbott, B. Derrida, M. Mezard, and J. J. Hopfield for discussions of many proposed solutions that would not work. The fertile intimate contact with biologists we owe to the Institute for Advanced Studies at the Hebrew University in Jerusalem. The work of D.J.A. was supported by a grant from the U.S.–Israel Binational Science Foundation.

1. Hopfield, J. J. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 2554–2558.
2. Morgestern, I. & Van Hemmen, L. B., eds. (1987) *Heidelberg Colloquium on Glassy Dynamics* (Springer, Heidelberg).
3. Amit, D. J. (1989) *Modeling Brain Function* (Cambridge Univ. Press, New York).
4. Kleinfeld, D. & Sompolinsky, H. (1988) *Biophys. J.* **54**, 1039–1051.
5. Amit, D. J. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 2141–2145.
6. Virasoro, M. A. (1988) *Europhys. Lett.* **7**, 293–298.
7. Amit, D. J., Sagi, D. & Usher, M. (1989) *Architecture of Attractor Neural Network Performing Cognitive Fast Scanning* (Racah Institute, Hebrew University, Jerusalem), preprint 89/306.
8. Amit, D. J., Gutfreund, H. & Sompolinsky, H. (1987) *Ann. Phys.* **173**, 30–67.
9. Amit, D. J., Gutfreund, H. & Sompolinsky, H. (1987) *Phys. Rev. A* **35**, 2293–2303.
10. Gardner, E. (1988) *J. Phys. A* **21**, 257–270.
11. Buhmann, J., Divko, R. & Schulten, K. (1988) in *Neural Networks From Models to Applications*, eds. Personnaz, L. & Dreyfus, G. (IDSET, Paris), pp. 360–371.
12. Tsodyks, M. V. & Feigl'man, M. V. (1988) *Europhys. Lett.* **6**, 101–105.
13. Shinomoto, S. (1987) *Biol. Cybern.* **57**, 197–206.
14. Willshaw, D. J., Buneman, O. P. & Longuet-Higgins, H. C. (1969) *Nature (London)* **222**, 960–961.
15. Anderson, R. A. & Mountcastle, V. B. (1983) *J. Neurosci.* **3**, 532–548.
16. Sur, M., Wall, J. T. & Kaas, J. H. (1984) *J. Neurophysiol.* **51**, 724–744.
17. Goldberg, M. E. & Bruce, C. J. (1985) *Vision Res.* **25**, 471–487.
18. Miyashita, Y. & Chang, H. S. (1988) *Nature (London)* **331**, 68–70.
19. Eccles, J. C. (1964) *Physiology of Synapses* (Springer, Berlin).
20. Fatt, P. & Katz, B. (1953) *J. Physiol. (London)* **121**, 374–389.
21. Glauber, R. J. (1963) *J. Math. Phys.* **4**, 294–307.
22. Amit, D. J., Gutfreund, H. & Sompolinsky, H. (1985) *Phys. Rev. A* **32**, 1007–1018.
23. Amit, D. J. & Treves, A. (1989) *J. Phys. A* **22**, 2205–2226.